



PRESENTACIÓN FINAL

Autores:

Ana María Soto

Yoselin Nieto

Nicolas Yepes

Felipe Lobato

Imagina que, en un festival de música, el escenario más popular es el de un género que muchos consideran 'poco comercial', pero que resuena profundamente con la audiencia. ¿Qué factores lo hacen atractivo?



PROBLEMA

¿Por qué algunos géneros musicales se vuelven más populares que otros en diferentes contextos económicos en EEUU?



CLIENTE POTENCIAL

- 🎵 En EEUU, hay alrededor de **20,000 a 30,000 productores musicales** registrados.
- 🎵 En la industria musical, el # de analistas de datos es más difícil de estimar, un estimado podría ser **5,000 a 10,000 analistas** trabajando directamente en el sector musical.
- 🎵 A nivel global, hay más de **300 plataformas de streaming** que ofrecen servicios de música, incluyendo las más populares como Spotify, Apple Music, Amazon Music, y Tidal



CONTEXTO DE SOLUCIÓN

La industria musical enfrenta retos al intentar predecir qué canciones resonarán con el público. La identificación de patrones puede ser clave.



Este proyecto utiliza aprendizaje no supervisado (Clustering) para explorar relaciones en los datos musicales y económicos sin etiquetas predefinidas, permitiendo descubrir patrones ocultos.

Clustering jerárquico

Kmeans

Kmedoides

DBSCAN



DATOS

1



Base descargado de
spotify

Variable	Tipo
artista	object
cancion	object
duracion_ms	int64
explicito	bool
anio	int64
popularidad	int64
bailable	float64
energia	float64
clave	int64
volumen	float64
modo	int64
locuacidad	float64
acusticidad	float64
instrumentalidad	float64
vivisidad	float64
valencia	
emocional	float64
tempo	float64
genero	object

■ Información original

■ Información incluida

2



Datos de la API spotify

Variable	Tipo
Numero de seguidores	int64
Fecha de lanzamiento completa	Datetime

3

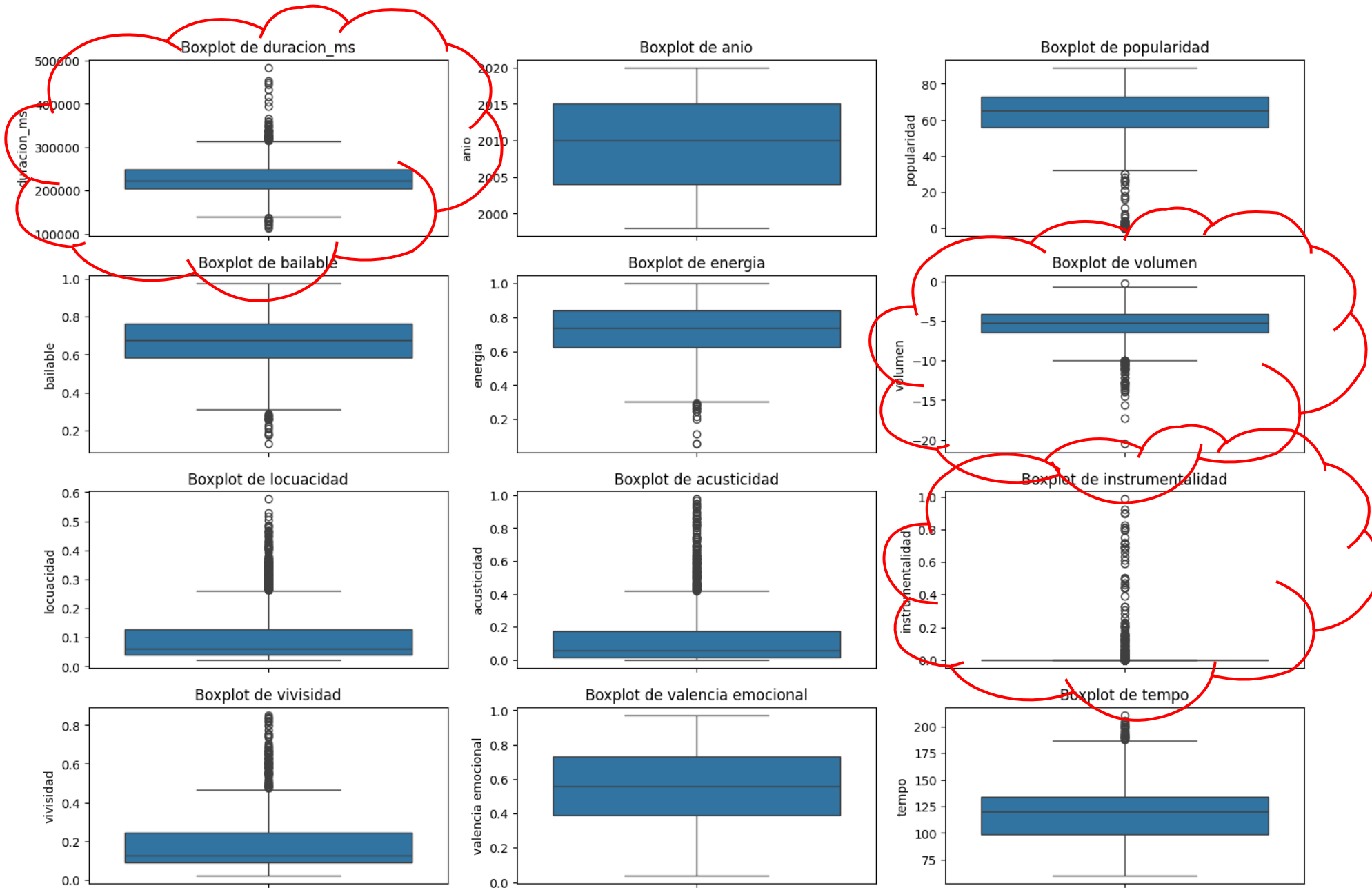


Datos Macroeconómicos de
EEUU

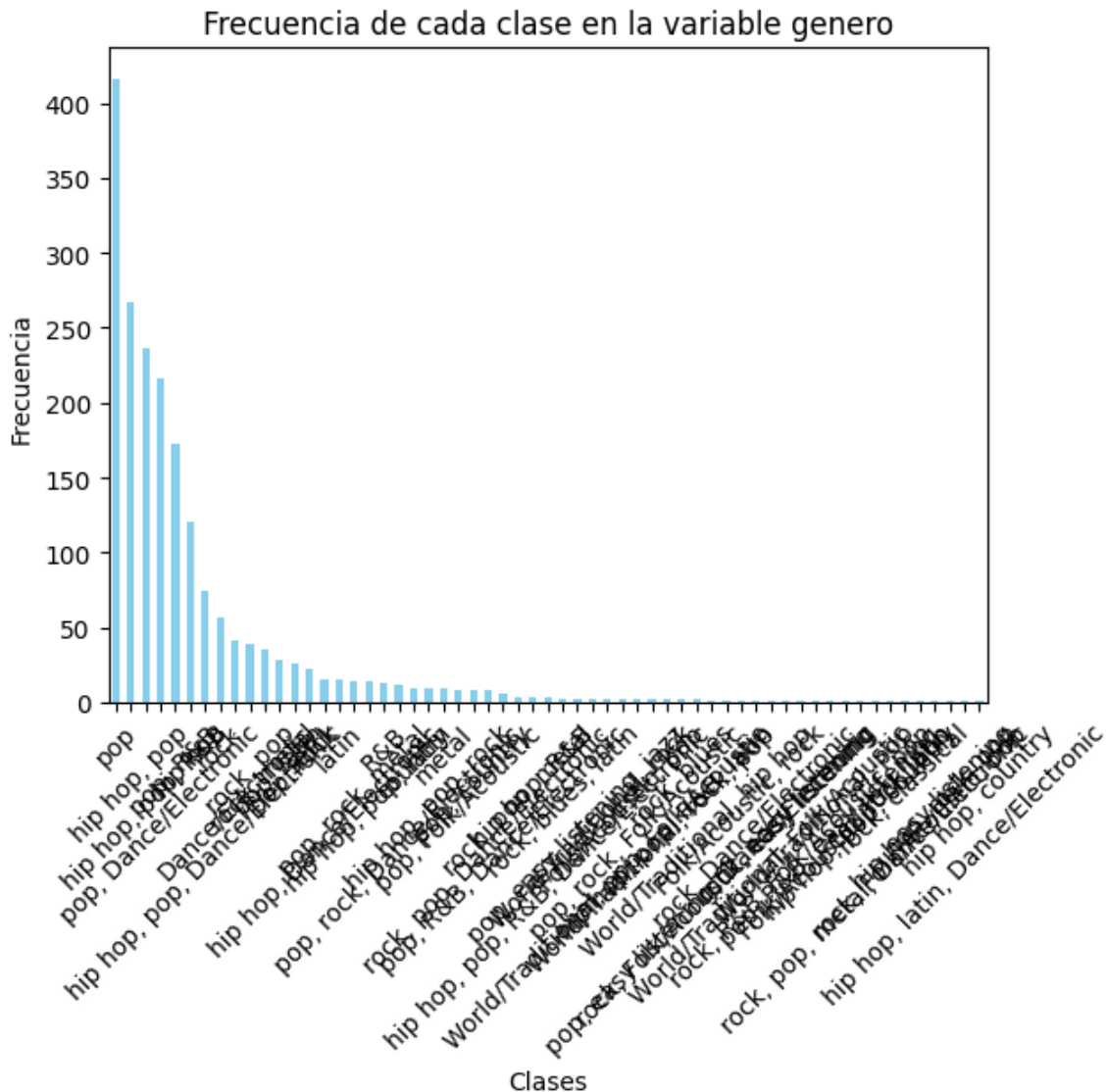
Variable	Tipo
inflation	float64
unemployment	float64
gdp_per_capita	float64
gross_savings	float64
crecimiento_renta_neta	float64
consumption_expenditure	float64

1940 REGISTROS APARTIR DEL AÑO 1999

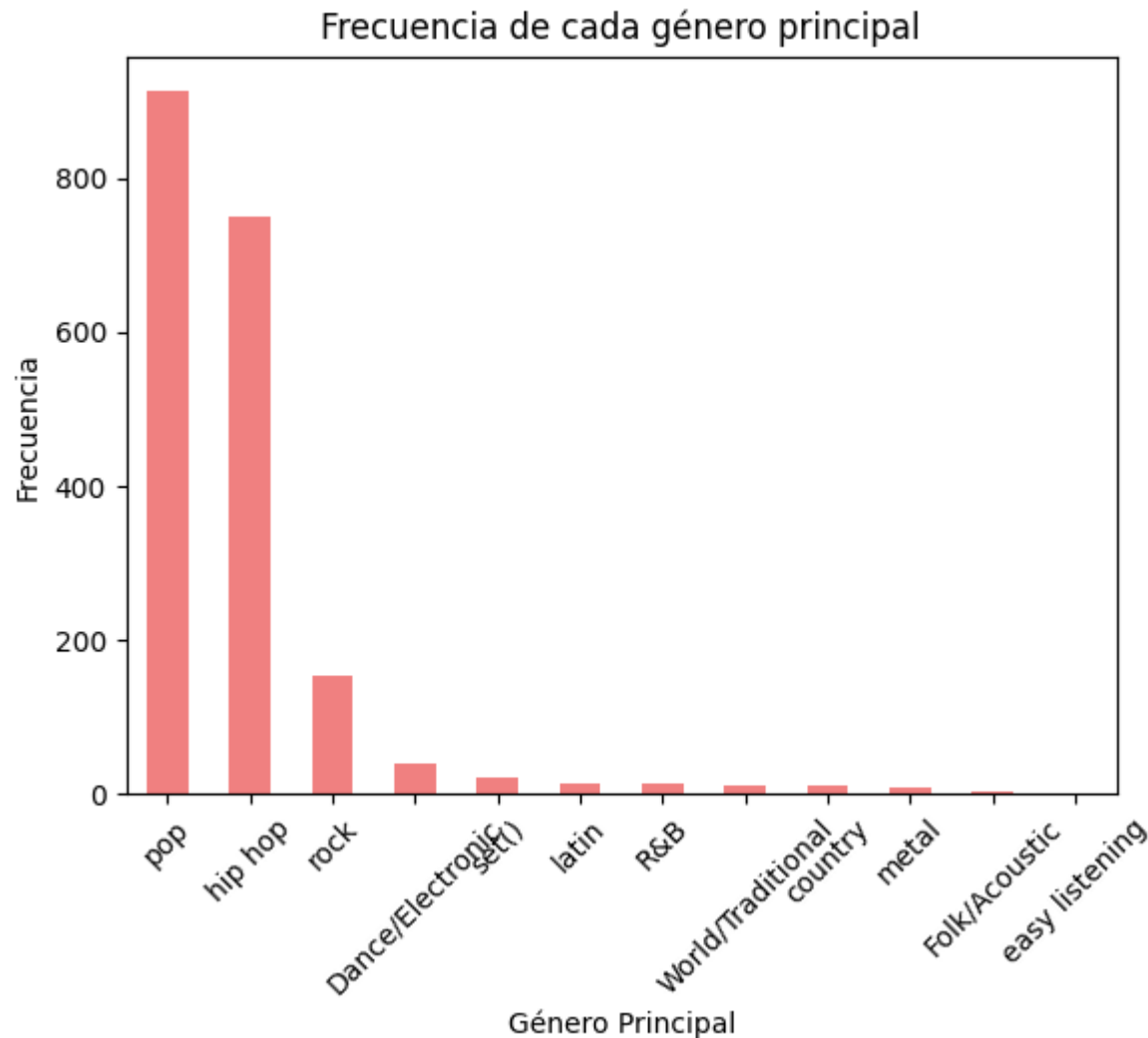
INFORMACIÓN ATIPICA DE LAS VARIABLES NÚMERICAS



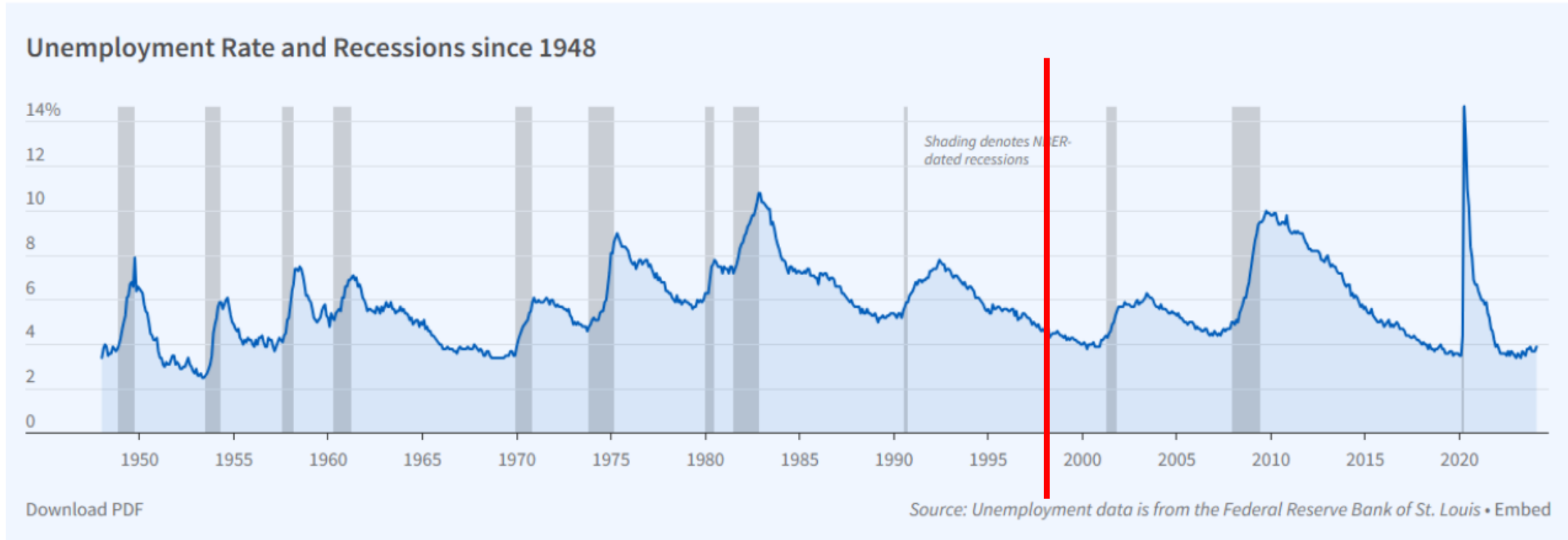
VARIABLE GENERO ORIGINAL



VARIABLE GENERO HOMOLOGADA

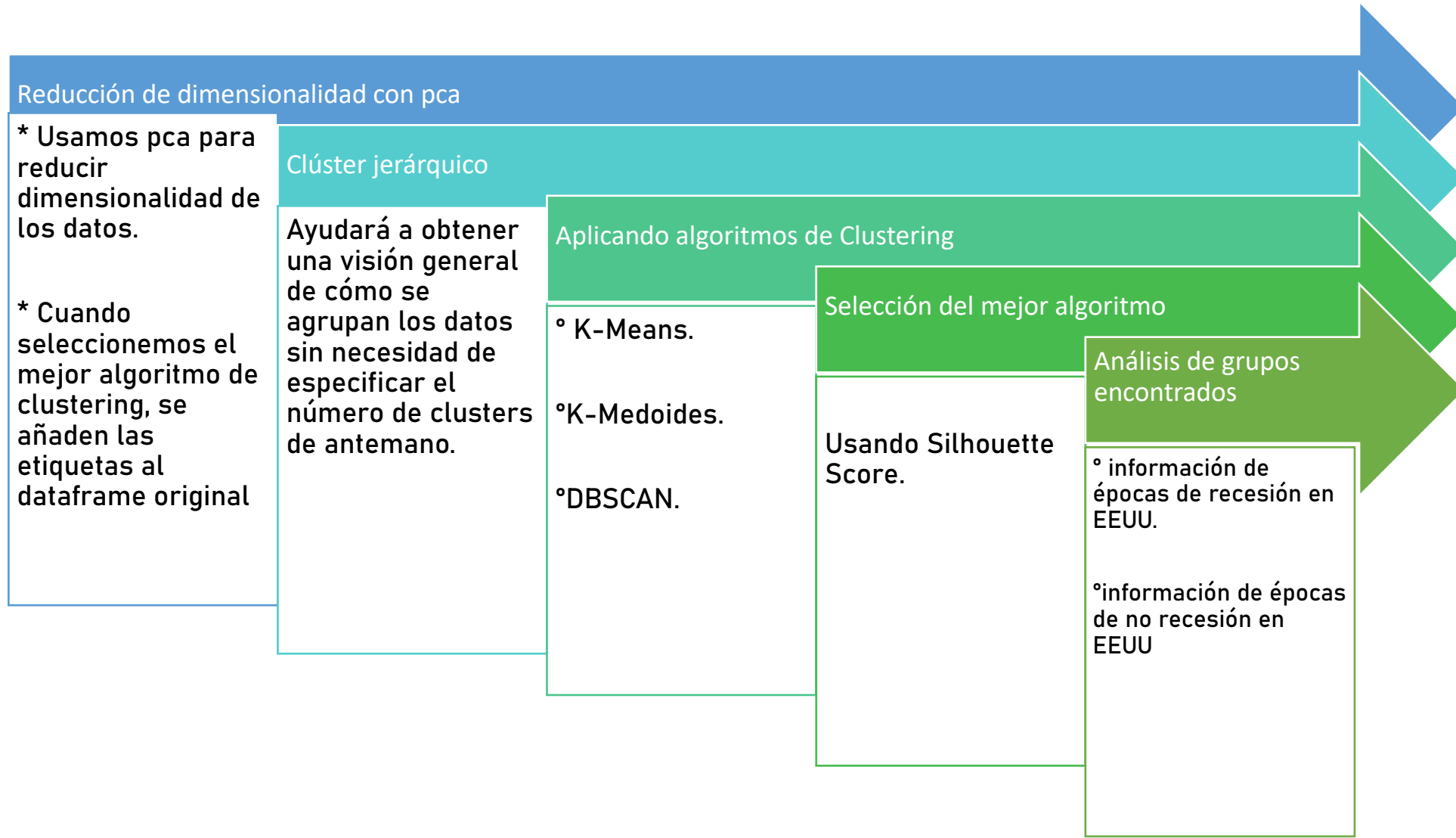


¿Cómo sectorizamos la información?



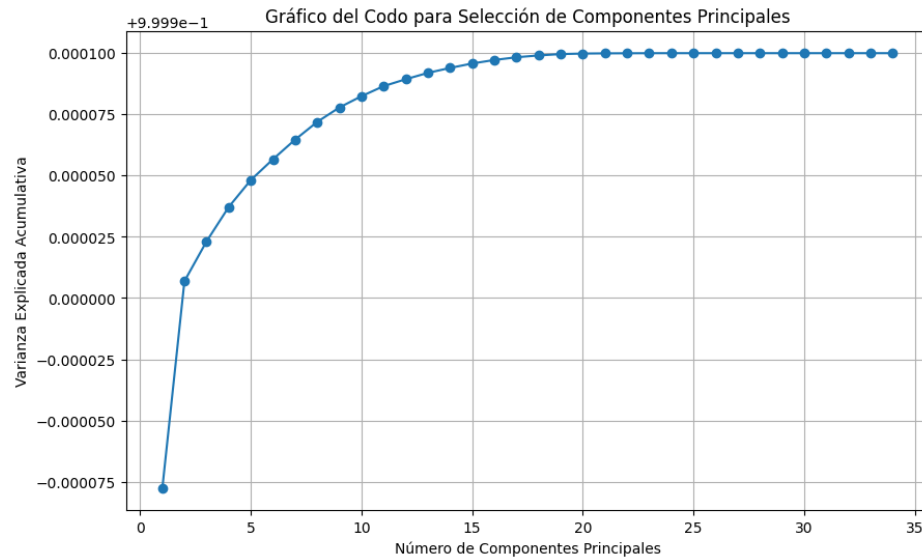
se creo una variable que identifico los periodos de bonanza según la fecha de lanzamiento de las canciones. Esta información se obtuvo a partir de los datos proporcionados por el Comité de Datación del Ciclo Económico de la NBER, que publica una cronología oficial de los ciclos económicos en Estados Unidos.

CLUSTERING

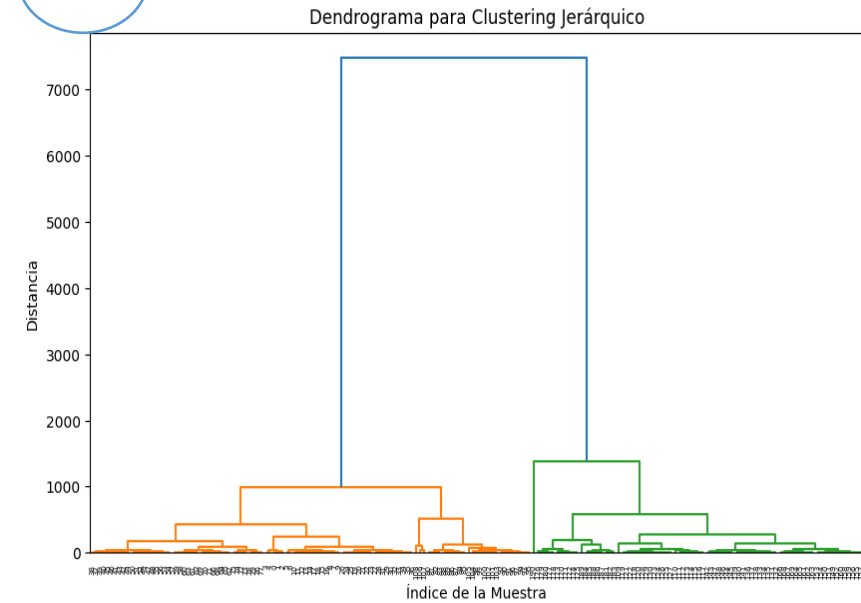


INFORMACIÓN DE RECESIÓN ECONOMICA

1



2



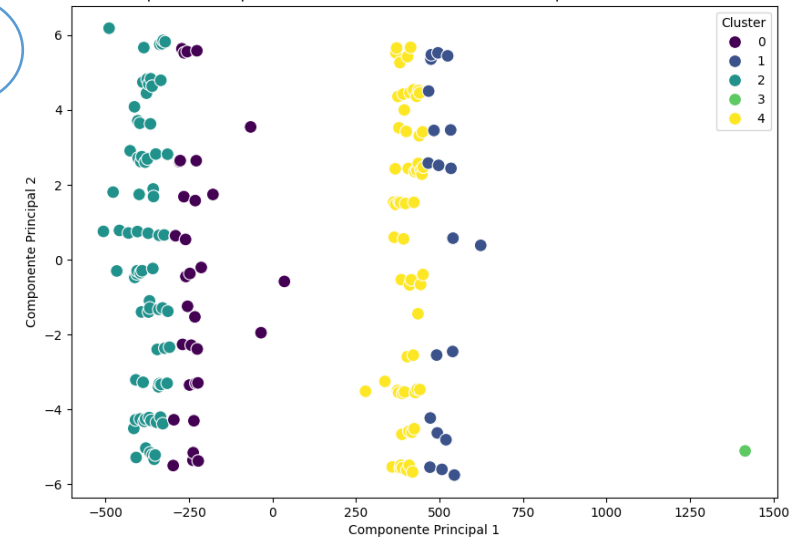
3

Comparado resultados

Silhouette	Scores:
KMeans:	0,61738
KMedoids:	0,5190
DBSCAN:	0,5944

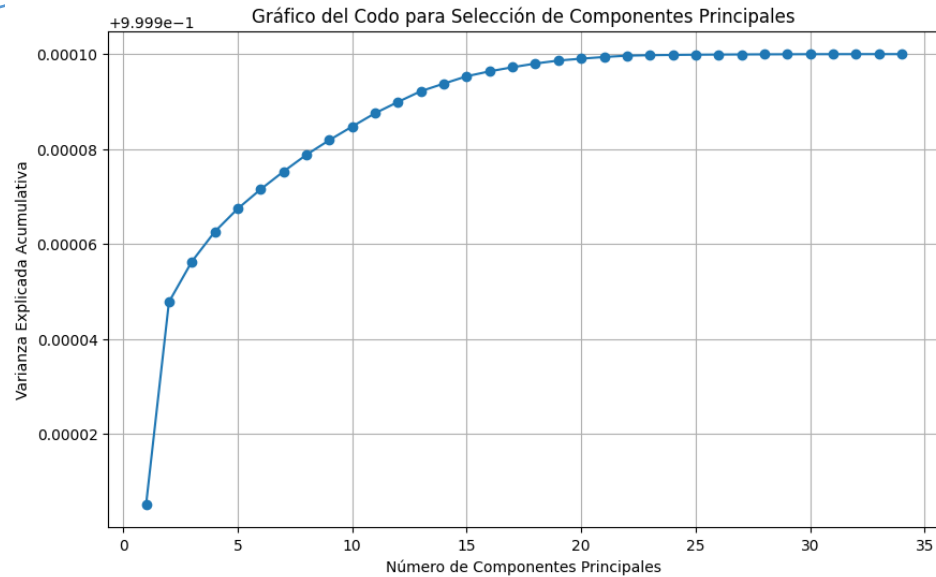
4

Clusters formados por KMeans para las características de la musica en Epoca de Recesiòn Economica en EEUU

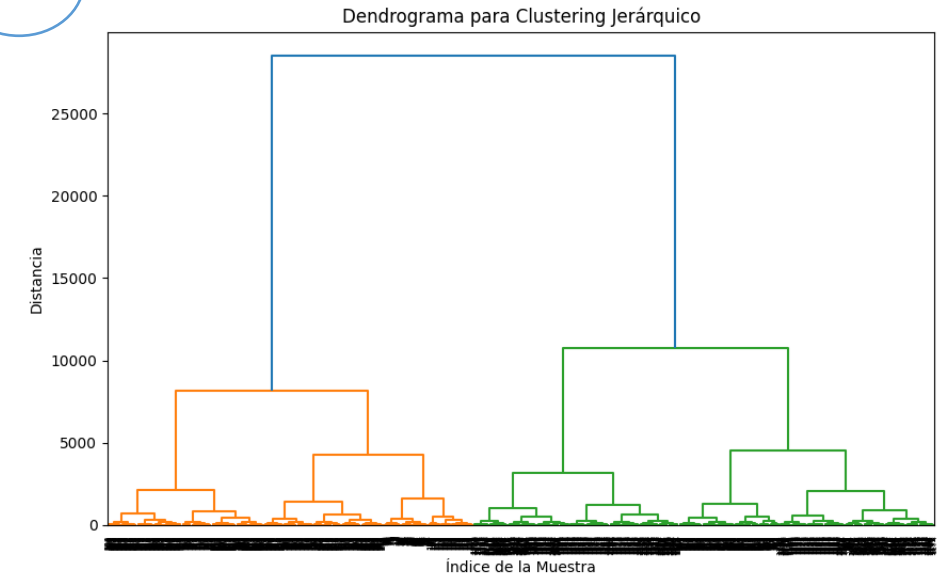


INFORMACIÓN DE NO RECESIÓN ECONOMICA

1



2



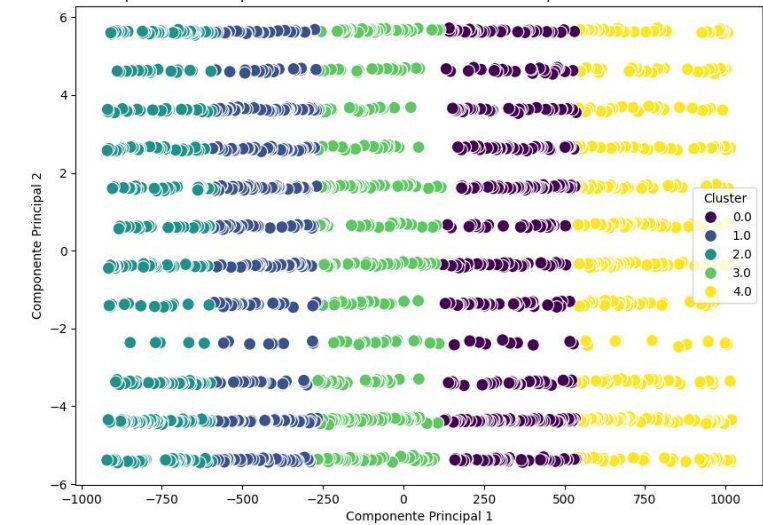
3

Comparado resultados

Silhouette	Scores:
KMeans:	0.5583714663
KMedoids:	0.5599593491
DBSCAN:	0.4099

4

Clusters formados por Kmedoides para las características de la musica en Epoca de No Recesión Economica en EEUU



CONCLUSIONES

- A pesar de que el hip hop, el pop y el rock tienen una prevalencia en la música que se escucha en periodos de recesión y bonanza (sea o no popular la música), si es posible encontrar características propias de cada uno de los clusters que indican que, además de la influencia cultural que determina los géneros de moda, la economía también tiene un papel que jugar para determinar cómo se escucha la música:
- **Cluster popular sin recesión económica:** Las canciones en este grupo tienen un alto índice de bailabilidad (0.318) y un tempo más alto respecto a otros clusters (0.06), con una locuacidad alta (0.21), baja instrumentalidad (-0.11), acústica de (0.26) y una valencia emocional negativa en sus temáticas.
- **Cluster impopular sin recesión económica:** su música se caracteriza por tener un contenido de bailabilidad significativamente bajo (0.088), alta instrumentalidad (0.12), baja acústica (0.04) y un tempo bajo (-0.013).
- **Cluster popular con recesión económica:** Se caracteriza por tener una duración de canciones especialmente baja, (0.18), un lenguaje explícito moderado (0.23), baja bailabilidad (-0.30), de energía relativamente alta (0.31), una muy baja instrumentalidad (0.084) y con una temática más negativa que la mayoría de clusters de 0.15. Llama la atención que la música de valencia emocional negativa suele ser más popular, independiente del contexto económico
- **Cluster impopular con recesión económica:** Estas se caracterizan por ser canciones más largas (0.26), más explícitas (0.33), un poco más bailables, aunque menos energéticas, y dos veces más positivas en la temática que manejan (0.33).

RECOMENDACIONES Y ESTUDIOS POSTERIORES

- Segmentar y especializar el análisis hacia géneros musicales más específicos, con el propósito de generar recomendaciones de características musicales hechas a la medida para discográficas y grupos musicales de un género musical en específico.
- Explorar el aplicar DBSCAN en lugar de kmedias revisando un amplio más amplio de los valores de min_samples, dado que este algoritmo es menos sensible frente a valores atípicos.
- Con el propósito de incluir un histórico más amplio de variables macroeconomías de periodo de recesión, se recomienda ampliar la base de datos musicales para canciones que hayan sido publicadas antes de 1999, limitante que se tuvo para el presente estudio.